



end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was constructed by Life Technologies. Contact: Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax: (1) 301 610 8371 Email: [liang@lifetech.com](mailto:liang@lifetech.com) URL: <http://fulllength.invitrogen.com>

BASE COUNT 192 a 262 c 251 g 170 t 2 others

Query Match 34.1%; Score 874.6; DB 9; Length 877;

Best Local Similarity 99.7%; Pred. No. 5.2e-126;

Matches 874; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

29 CAGAGACCCCAAGCCTGTGCTGTCGACAGAGCTCAAAAGGCGCTTCCCTCC 88  
 1 CAGAGACCCCAAGCCTGTGCTGTCGACAGAGCTCAAAAGGCGCTTCCCTCC 60  
 89 TGGCTCGGCTGTGCTGTCGACAGAGCTCAAAAGGCGCTTCCCTCC 148  
 61 TGGCTCGGCTGTGCTGTCGACAGAGCTCAAAAGGCGCTTCCCTCC 120  
 149 TGATCCATCCCTGTGTCGACAGAGCTCAAAAGGCGCTTCCCTCC 208  
 121 TGATCCATCCCTGTGTCGACAGAGCTCAAAAGGCGCTTCCCTCC 180  
 209 CACCTAGCCTCTCCCTGAGAGCTCAAAAGGCGCTTCCCTCC 268  
 181 CACCTAGCCTCTCCCTGAGAGCTCAAAAGGCGCTTCCCTCC 240  
 269 AGGACCTGGGAGAGCTCAAAAGGCGCTTCCCTCC 328  
 241 AGGACCTGGGAGAGCTCAAAAGGCGCTTCCCTCC 300  
 329 CGTCTCAGAGAGCTGTCTTCCCAAGCCTTTGATGACAAACATTTCCCTGATGAT 388  
 301 CGTCTCAGAGAGCTGTCTTCCCAAGCCTTTGATGACAAACATTTCCCTGATGAT 360  
 389 GCTTCTGAGTGTCTGTCGACAGAGCTCAAAAGGCGCTTCCCTCC 448  
 361 GCTTCTGAGTGTCTGTCGACAGAGCTCAAAAGGCGCTTCCCTCC 420  
 449 CAGCCCAAGCTTGAAGTGTCTGTCGACAGAGCTCAAAAGGCGCTTCCCTCC 508  
 421 CAGCCCAAGCTTGAAGTGTCTGTCGACAGAGCTCAAAAGGCGCTTCCCTCC 480  
 509 GAGCAAGGCGCAAGCGGTGCGTGGGAGAGTTCGCGGAGGTGCGCGGCGAGCTGT 568  
 481 GAGCAAGGCGCAAGCGGTGCGTGGGAGAGTTCGCGGAGGTGCGCGGCGAGCTGT 540  
 569 CGCTGAGAGCTCGGGAGAGCAATTGACCTGTCTGAGAGTGAAGCTGTGTCG 628  
 541 CGCTGAGAGCTCGGGAGAGCAATTGACCTGTCTGAGAGTGAAGCTGTGTCG 600  
 629 TGTCTGAAGCTCAAGGAGAGATTAATATCCCAAGGTCACAGGCGCAAAATCTCCC 688  
 601 TGTCTGAAGCTCAAGGAGAGATTAATATCCCAAGGTCACAGGCGCAAAATCTCCC 660  
 689 ATGGGTGCTGATAGAGGCTCTGATCGGGAGAGCAAGAGAACTGCTGTGTACTCTG 748  
 661 ATGGGTGCTGATAGAGGCTCTGATCGGGAGAGCAAGAGAACTGCTGTGTACTCTG 720  
 749 GGAACCTGGAAGGCGCTTCTGATCGGGAGAGCAAGAGAACTGCTGTGTACTCTG 808  
 721 GGAACCTGGAAGGCGCTTCTGATCGGGAGAGCAAGAGAACTGCTGTGTACTCTG 780  
 809 TGTCAATCGGCTCAAGCGGCTGATCTCTGAGAGCTCAAAAGGCGCTTCCCTCC 868  
 781 TGTCAATCGGCTCAAGCGGCTGATCTCTGAGAGCTCAAAAGGCGCTTCCCTCC 840  
 869 GCGTGAAGAGGCTGTGATATCTCAAGCGGCT 905

Db 841 GCCTTGACATGCGCTGTGATCTACAGGCGCT 877

RESULT 2  
 B0054281  
 LOCUS  
 DEFINITION  
 AGNCOURT 6830234 NIH\_MGC\_106 Homo sapiens cDNA clone IMAGE:5936362  
 5' mRNA sequence.

ACCESSION  
 B0054281  
 VERSION  
 B0054281.1 GI:19813621

KEYWORDS  
 SOURCE  
 ORGANISM

Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE  
 NIH-MGC <http://mgc.nci.nih.gov/>  
 1 (bases 1 to 1020)

ATTORNS  
 TITLE  
 JOURNAL  
 COMMENT

Contact: Robert Strausberg, Ph.D.  
 Email: [cgabs-remail.nih.gov](mailto:cgabs-remail.nih.gov)  
 Tissue Procurement: Dr. Daniel McVicar, DMS/MCI  
 cDNA Library Preparation: Rubin Laboratory  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LINL at:  
<http://image.llnl.gov>  
 Plate: LICM2125 row: j column: 11  
 High quality sequence stop: 555.  
 Location/Qualifiers

FEATURES  
 source

1..1020  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone\_id="IMAGE:5936362"  
 /clone\_lib="NIH\_MGC\_106"  
 /tissue\_type="natural killer cells, cell line"  
 /lab\_host="DH10B (phage-resistant)"  
 /note="Organ: blood; Vector: pOT8; Site 1: XhoI; Site 2:  
 EcoRI; cDNA made by oligo-dT priming. Directionally cloned  
 into EcoRI/XhoI sites using the following 5' adaptor:  
 GGCAGAG(G). Library constructed by Ling Hong in the  
 Laboratory of Gerald M. Rubin (University of California,  
 Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and  
 Superscript II RT (Life Technologies). Note: this is a  
 NIH MGC Library."

BASE COUNT 219 a 311 c 283 g 204 t 3 others

Query Match 33.6%; Score 862.6; DB 14; Length 1020;

Best Local Similarity 95.5%; Pred. No. 3.4e-124;

Matches 973; Conservative 0; Mismatches 36; Indels 10; Gaps 8;

204 ACCAACAACCTAGCCTCTCCCTGAAGATCTCCAGAGCTGAGAGATTTGGGTCTAG 263  
 1 AACACACATAGCCTCTCCCTGAAGATCTCCAGAGCTGAGAGATTTGGGTCTAG 60  
 264 GACCAAGAGCACTGGGAGAGCTTCCAGAAAGGCGCCCAAGCCCTTAACCTGACGCGA 323  
 61 GACCAAGAGCACTGGGAGAGCTTCCAGAAAGGCGCCCAAGCCCTTAACCTGACGCGA 120  
 324 GATGCTGCTGAGAGCTGTCTTCCCAAGCTTTGATGACAAACATTTCCCTGAT 383  
 121 GATGCTGCTGAGAGCTGTCTTCCCAAGCTTTGATGACAAACATTTCCCTGAT 180  
 384 GATGCTGCTGAGAGCTGTCTTCCCAAGCTTTGATGACAAACATTTCCCTGAT 443  
 181 GATGCTGCTGAGAGCTGTCTTCCCAAGCTTTGATGACAAACATTTCCCTGAT 240  
 444 TCTGCAAGCCCAAGCTTGAAGTTCCTGTCTCAAGGCGAGGAGCTGTGACATGGAAC 503  
 241 TCTGCAAGCCCAAGCTTGAAGTTCCTGTCTCAAGGCGAGGAGCTGTGACATGGAAC 300



|    |      |  |      |
|----|------|--|------|
| Qy | 1039 | ACTGTGCAGAGACAACACTCACTGGAAAAGGCTGACAGCTCCCTCTGTTTTCTGAA     | 1098 |
| Db | 720  | ACTGTGCAGAGACAACACTCACTGGAAAAGGCTGACAGCTCCCTCTGTTTTCTGAA     | 779  |
| Qy | 1099 | GCTGTGACA-GGGAGAGACTCTTTCTCAGTAAAGG--TCTCCGGAGTCCCTCAGCTTC   | 1155 |
| Db | 780  | ACTGTGCAGAGGGGAGAAATCTTTCTCAGTAAAGGCTCTCCGGAGTCCCTCAGCTTC    | 839  |
| Qy | 1156 | TACATCAGCCGGAATGACA-GGCTGTCTCTTTGGATGATG-CTTAGGGCCAAAGAGAG   | 1213 |
| Db | 840  | TACATCAGCCCTGATGACGAGGCTGTCTCTTTGGATGATGATGCCCTTAGCCCAAGAGAG | 899  |
| Qy | 1214 | GCCAAAAGGG--AAACAAGGCTGCACACT-AGAACCCCAATTACGCT--CCTGGGA     | 1268 |
| Db | 900  | GCCAAAAGGAAAAACAGGGCTGCACACTTAAAAACCCCAATTTAGCTTCTCTTGGCC    | 959  |
| Qy | 1269 | CCCCGAGGCAAGCTGTGCACACTAGGAGAGGAGG                           | 1304 |
| Db | 960  | CCCCAAAAGCACTGAGGCTCTTCCGGGAAAAAG                            | 995  |

|            |  |        |      |                        |
|------------|--|--------|------|------------------------|
| RESULT 4   |  |        |      |                        |
| LOCUS      | AL568702/c   |        |      |                        |
| DEFINITION | AL568702 LTI_FLO02.PL1 Homo sapiens cDNA clone CSDBE005TK23 3 prime  | 854 bp | mRNA | linear EST 16-FEB-2001 |
| ACCESSION  | ' mRNA sequence. _.. '   |        |      |                        |
| VERSION    | AL568702   |        |      |                        |
| KEYWORDS   | AL568702.1 GI:12923305   |        |      |                        |
| SOURCE     | EST.   |        |      |                        |
| ORGANISM   | human.   |        |      |                        |
|            | Homo sapiens   |        |      |                        |
|            | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  |        |      |                        |
|            | Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  |        |      |                        |
| REFERENCE  | 1 (bases 1 to 854)   |        |      |                        |
| AUTHORS    | Li,W.B., Gruber,C., Jessee,J. and Polayes,D.   |        |      |                        |
| TITLE      | Full-length cDNA libraries and normalization   |        |      |                        |
| JOURNAL    | Unpublished (2001)   |        |      |                        |
| COMMENT    | Contact: Genoscope<br>Genoscope - Centre National de Sequencage<br>BP 191 91006 EVRY cedex - France<br>Email: sege@genoscope.cns.fr, Web : www.genoscope.cns.fr. |        |      |                        |
| FEATURES   | Location/Qualifiers  |        |      |                        |

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="SC0DE005YK23"
/clone_1id="U11_F1002_P11"
/lab_host="DH10B"
/notes="Organ: Placenta; Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville , Maryland 20850, USA Fax : (1) 301 610 8371 Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com"

```



QY 53 GTGACGAGCTCAAGGGCCCTGGCCCTCCCTGCTGCTGCTGAGGCT 112  
 Db 1 GTGACGAGCTCAAGGGCCCTGGCCCTCCCTGCTGCTGCTGAGGCT 59  
 QY 113 TCCCGAGTCCGAATTCCTTAAGAGAGAGAGAGAGAGAGAGAGAG 172  
 Db 60 TCCCGAGTCCGAATTCCTTAAGAGAGAGAGAGAGAGAGAGAGAG 119  
 QY 173 GCTGACTGACAGATGCTGAGTACCAACCAACCAACCAACCAACCA 232  
 Db 120 GCTGACTGACAGATGCTGAGTACCAACCAACCAACCAACCAACCA 179  
 QY 233 TCCGAGCTGAGAGAGTCTGAGTCTGAGAGAGAGAGAGAGAGAGAG 292  
 Db 180 TCCGAGCTGAGAGAGTCTGAGTCTGAGAGAGAGAGAGAGAGAGAG 239  
 QY 293 GGGCCCCAAGCCCTTAACCTGTCACAGCCAGAGAGAGAGAGAGAG 352  
 Db 240 GGGCCCCAAGCCCTTAACCTGTCACAGCCAGAGAGAGAGAGAGAG 299  
 QY 353 AGCCTTGATGACAAACCAATTCCTGATGATGATGATGATGATGAT 412  
 Db 300 AGCCTTGATGACAAACCAATTCCTGATGATGATGATGATGATGAT 359  
 QY 413 CAATGGGAAGTCTGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 472  
 Db 360 CAATGGGAAGTCTGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 419  
 QY 473 TCCGAGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 532  
 Db 420 TCCGAGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 478  
 QY 533 TGGGCAAGTTCCCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 592  
 Db 479 TGGGCAAGTTCCCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 538  
 QY 593 CCATGCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 652  
 Db 539 CCATGCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 598  
 QY 653 ATAACTCCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 712  
 Db 599 ATAACTCCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 658  
 QY 713 GCAGGAG 769  
 Db 659 GCAGGAG 718  
 QY 770 TCATCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 803  
 Db 719 CATCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 752  
 RESULT 7  
 AL844325 677 bp mRNA linear EST 30-JUL-2002  
 DEFINITION AL844325 pool\_YT\_11b\_v\_SPD Homo sapiens cDNA, mRNA sequence.  
 ACCESSION AL844325  
 VERSION AL844325.1 GI:22019105  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
 1 (bases 1 to 677)  
 AUTHORS Ashcroft, K., Bethel, G., Bye, J.M., Howell, G.R., Huckle, E.J. and  
 Sheridan, E.  
 TITLE Homo sapiens EST sequence  
 JOURNAL Unpublished (2002)  
 COMMENT Contact: The Sanger Centre  
 The Sanger Centre  
 Wellcome Trust Genome Campus, Hinxton, Cambridgeshire, CB10 1SA, UK

Email: humquery@sanger.ac.uk  
 Sanger Centre name: scsdl0827.400489S  
 Homo sapiens EST sequence. This sequence was generated as part of  
 The Wellcome Trust Sanger Institute program to identify and  
 annotate genes in the human genome. Incomplete or unconfirmed genes  
 are experimentally analysed using a variety of cDNA library  
 resources. This sequence was obtained from a PCR product generated  
 from a pool of up to 100,000 cDNA clones derived from  
 pool\_YT\_11b\_v\_SPD cDNA library. Further information can be found at  
 http://www.sanger.ac.uk/teams/team69/.  
 Location/Qualifiers  
 1..677  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /map="20"  
 /clone\_11b="pool\_YT\_11b\_v\_SPD"  
 /note="Organ: breast; Vector: pZero-1; Site 1: SpH1;  
 Site 2: SpH1; Ductal carcinoma in situ, high-grade, comedo  
 , from 41 yo female. Library constructed in the laboratory  
 of K. Polyak, M.D., Ph.D. (Dana-Farber Cancer Institute)."  
 BASH COUNT 160 a 195 c 154 g 168 t  
 ORIGIN  
 Query Match 25 94; Score 666; DB 9; Length 677;  
 Best Local Similarity 99.9%; Pred. No. 1e-93; 0; Indels 1; Gaps 1;  
 Matches 677; Conservative 0; Mismatches 0;  
 QY 1680 TGGCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1739  
 Db 1 TGGCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 60  
 QY 1740 GTCCAGGCTCATGATGATGATGATGATGATGATGATGATGATGAT 1799  
 Db 61 GTCCAGGCTCATGATGATGATGATGATGATGATGATGATGATGAT 120  
 QY 1800 AATGTCCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1859  
 Db 121 AATGTCCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 179  
 QY 1860 TGGCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1919  
 Db 180 TGGCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 239  
 QY 1920 CCTGATTCGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1979  
 Db 240 CCTGATTCGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 299  
 QY 1980 CCACCATGCTGGCTAATTTTGTATTTTGTATTTTGTATTTTGTATTT 2039  
 Db 300 CCACCATGCTGGCTAATTTTGTATTTTGTATTTTGTATTTTGTATTT 359  
 QY 2040 CAGGCTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2039  
 Db 360 CAGGCTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 419  
 QY 2100 GGGATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2159  
 Db 420 GGGATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 479  
 QY 2160 GGGATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2219  
 Db 480 GGGATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 539  
 QY 2220 TCTTAAGCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2279  
 Db 540 TCTTAAGCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 599  
 QY 2280 AGGTGCTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2339  
 Db 600 AGGTGCTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 659  
 QY 2340 CCGTTGATTCCTGCTG 2357

Db 660 CCCTTGATTCCCTGCTG 677

# RESULT 8

LOCUS B0054265 986 bp mRNA linear EST 29-MAR-2002  
DEFINITION AGENCOURT 6830248 NIH\_MGC\_106 Homo sapiens cDNA clone IMAGE:5936339  
5', mRNA sequence.

ACCESSION B0054265  
VERSION B0054265.1 GI:19813605  
KEYWORDS EST.

## SOURCE

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

1 (bases 1 to 986)  
NIH-MGC http://mgs.nci.nih.gov/

National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)

CONTACT: Robert Strausberg, Ph.D.  
Email: cgapbs-remail.nih.gov

Tissue Procurement: Dr. Daniel McVicar, DSI/NCI  
CDNA Library Preparation: Rubin Laboratory

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNL at:

http://image.lnl.gov  
Plate: LLM2125 row: 1 column: 12

High quality sequence stop: 515.  
Location/Qualifiers

## FEATURES

source

1..986

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone\_image="5936339"

/clone\_id="NIH\_MGC\_106"

/tissue\_type="natural killer cells, cell line"

/lab\_host="DH10B (phage-resistant)"

/note="Organ: blood; Vector: pOT7; Site 1: XhoI; Site 2:  
EcoRI; CDNA made by oligo-dT priming. Directionally cloned  
into EcoRI/XhoI sites using the following 5' adaptor:  
GGCAGGAG(G) library constructed by Ling Hong in the  
laboratory of Gerald M. Rubin (University of California,  
Berkeley) using ZAP-CDNA synthesis kit (Stratagene) and  
Superscript II RT (Life Technologies). Note: this is a  
NIH MGC Library."

BASE COUNT 211 a 291 c 276 g 207 t 1 others

## ORIGIN

Query Match 25.9%; Score 664.2; DB 14; Length 986;  
Best Local Similarity 98.8%; Pred. No. 1.5e-93;

Matches 722; Conservative 0; Mismatches 3; Indels 6; Gaps 5;

137 GCATGGGAGCTGATCCATCCCTGGTGTGCAAACTGCTGACAGAGAGTCTGAGC 196

1 GCATGGGAGCTGATCCATCCCTGGTGTGCAAACTGCTGACAGAGAGTCTGAGC 60

197 TACCCAAACCAACCTAGCTCTCCCTGAAGATCTCCGAGCTGAGAGATCTGAGT 256

61 TACCCAAACCAACCTAGCTCTCCCTGAAGATCTCCGAGCTGAGAGATCTGAGT 120

257 GTCTTAGAGACCAAGACACTGGGAGACTTCCAGAAAGGCCCCCAAAAGCCTTA 316

121 GTCTTAGAGACCAAGACACTGGGAGACTTCCAGAAAGGCCCCCAAAAGCCTTA 180

317 AGCCAGAGCATGCTCTCAGAGAGCTGTCTCCCAAGCCTTTGATGAGCAAACTTTC 376

181 AGCCAGAGCATGCTCTCAGAGAGCTGTCTCCCAAGCCTTTGATGAGCAAACTTTC 240

377 CCTGATGATGCTCTTCTGAGAGTCTCTGCTGAGAGAAAGTGGAGAGTCTGAG 436

241 CCTGATGATGCTCTTCTGAGAGTCTCTGCTGAGAGAAAGTGGAGAGTCTGAG 300

QY 437 GAAATCTGCGCAAGCCCAAGCTTGAGTTCTCTGTCCAGAGCCAGGAGCTGTGACCA 496

Db 301 GAAATCTGCGCAAGCCCAAGCTTGAGTTCTCTGTCCAGAGCCAGGAGCTGTGACCA 360

QY 497 TGAAG 556

Db 361 TGAAG 420

QY 557 CGGCGAGAGCTGTCTGAGAGACTCGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 616

Db 421 CGGCGAGAGCTGTCTGAGAGACTCGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 480

QY 617 GGTGAG 676

Db 481 GGTGAG 540

QY 677 CCAAGCTTCCAG 736

Db 541 CCAAGCTTCCAG 600

QY 737 TGTGTACCTGAG 794

Db 601 TGTGTACCTGAG 660

QY 795 AGGCTTCTACTCTCTCTGAGT-CCGCTCAGAGAGAGAGAGAGAGAGAGAGAGAG 850

Db 661 AGGCTTCTACTCTCTCTGAGT-CCGCTCAGAGAGAGAGAGAGAGAGAGAGAGAG 720

QY 851 GACACTACAGAG 861

Db 721 GACACTACAGAG 731

## RESULT 9

LOCUS BG677567 794 bp mRNA linear EST 01-MAY-2001

DEFINITION 60262411891 NCI\_CGAP\_Skn4 Homo sapiens cDNA clone IMAGE:4748884 5',  
mRNA sequence.

ACCESSION BG677567  
VERSION BG677567.1 GI:13908964  
KEYWORDS EST.

## SOURCE

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

1 (bases 1 to 794)  
NIH-MGC http://mgs.nci.nih.gov/

National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)

CONTACT: Robert Strausberg, Ph.D.  
Email: cgapbs-remail.nih.gov

Tissue Procurement: James Cleaver, M.D.  
CDNA Library Preparation: Life Technologies, Inc.

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNL at:

http://image.lnl.gov  
Plate: LLM10601 row: d column: 05

High quality sequence stop: 790.  
Location/Qualifiers

## FEATURES

source

1..794

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone\_image="4748884"

/clone\_id="NCI\_CGAP\_Skn4"

/tissue\_type="squamous cell carcinoma"

/lab\_host="DH10B (T1 phage-resistant)"

/note="Organ: skin; Vector: pCMV-SPORT6; Site 1: NotI;  
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.  
Average insert size 1.5kb. Library constructed by Life  
Technologies. Note: this is a NCI\_CGAP Library."

BASE COUNT 192 a 229 c 212 g 161 t

## ORIGIN

Query Match 25.5%; Score 653.8; DB 12; Length 794;  
 Best Local Similarity 98.6%; Pred. No. 7, 2e-92;  
 Matches 681; Conservative 0; Mismatches 7; Indels 3; Gaps 2;

Qy 930 GGTGACATTACTCTGAGTGGCGGATGACATCTGCTCCCTCACTCAAGAGGCGCTGT 989  
 Db 1 GGAGGTCTCTCTCAAGAGTGGCGGATGACATCTGCTCCCTCACTCAAGAGGCGCTGT 60  
 Qy 990 CCTGAGAGGCGTGGCGCGCTCCCTGAGAGATATACCCCTACCTGTGACTGTGAGAG 1049  
 Db 61 CCGCAGAGGCGTGGCGCGCTCCCTGAGAGATATACCCCTACCTGTGACTGTGAGAG 120  
 Qy 1050 GACACCACTCACTGAGAGAGAGTGGAGAGTGGCGGATGACATCTGCTCCCTCACTCAAGAGGCGCTGT 1109  
 Db 121 GACACCACTCACTGAGAGAGAGTGGAGAGTGGCGGATGACATCTGCTCCCTCACTCAAGAGGCGCTGT 180  
 Qy 1110 GGAGAGTCTCTCTCACTGAGAGAGTGGAGAGTGGCGGATGACATCTGCTCCCTCACTCAAGAGGCGCTGT 1169  
 Db 181 GGAGAGTCTCTCTCACTGAGAGAGTGGAGAGTGGCGGATGACATCTGCTCCCTCACTCAAGAGGCGCTGT 240  
 Qy 1170 TGACAGAGTCTCTCTCTGAGAGAGTGGAGAGTGGCGGATGACATCTGCTCCCTCACTCAAGAGGCGCTGT 1229  
 Db 241 TGACAGAGTCTCTCTCTGAGAGAGTGGAGAGTGGCGGATGACATCTGCTCCCTCACTCAAGAGGCGCTGT 300  
 Qy 1230 AGGCTGACACCTAGAACCCCAATTCAGCTCTGAGAGAGTGGAGAGTGGCGGATGACATCTGCTCCCTCACTCAAGAGGCGCTGT 1289  
 Db 301 AGGCTGACACCTAGAACCCCAATTCAGCTCTGAGAGAGTGGAGAGTGGCGGATGACATCTGCTCCCTCACTCAAGAGGCGCTGT 360  
 Qy 1290 CTCAGAGAGAGAGTGGAGAGTGGAGAGTGGCGGATGACATCTGCTCCCTCACTCAAGAGGCGCTGT 1349  
 Db 361 CTCAGAGAGAGAGTGGAGAGTGGAGAGTGGCGGATGACATCTGCTCCCTCACTCAAGAGGCGCTGT 420  
 Qy 1350 TTCCTCTTAAAGCTTAAAGTCACTTCTTCAAGTGCATGATCCCACTGCA 1409  
 Db 421 TTCCTCTTAAAGCTTAAAGTCACTTCTTCAAGTGCATGATCCCACTGCA 480  
 Qy 1410 CCTCTAGTGCAGTGCAGAGAGTGGAGAGTGGCGGATGACATCTGCTCCCTCACTCAAGAGGCGCTGT 1469  
 Db 481 CCTCTAGTGCAGTGCAGAGAGTGGAGAGTGGCGGATGACATCTGCTCCCTCACTCAAGAGGCGCTGT 540  
 Qy 1470 CTCCTGAGAGAGTGGAGAGTGGAGAGTGGCGGATGACATCTGCTCCCTCACTCAAGAGGCGCTGT 1529  
 Db 541 CTCCTGAGAGAGTGGAGAGTGGAGAGTGGCGGATGACATCTGCTCCCTCACTCAAGAGGCGCTGT 600  
 Qy 1530 CCTGAGAGAGTGGAGAGTGGAGAGTGGCGGATGACATCTGCTCCCTCACTCAAGAGGCGCTGT 1589  
 Db 601 CCTGAGAGAGTGGAGAGTGGAGAGTGGCGGATGACATCTGCTCCCTCACTCAAGAGGCGCTGT 658  
 Qy 1590 GAAACAACTCTAGATCAACGAGAAATCA 1620  
 Db 659 GAAACAA-CCCTAGATCAACGAGAAATCA 688

RESULT 10  
 AL844308/c 642 bp mRNA linear EST 30-JUL-2002  
 LOCUS AL844308  
 DEFINITION AL844308 pool\_AK\_11b\_v\_SPD Homo sapiens cDNA, mRNA sequence.  
 ACCESSION AL844308  
 VERSION AL844308.1 GI:22019090  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 642)  
 AUTHORS Aheric, K., Bechel, G., Bye, J.M., Howell, G.R., Huckle, E.J. and  
 Sheridan, E.  
 TITLE Homo sapiens EST sequence  
 JOURNAL Unpublished (2002)  
 COMMENT Contact: The Sanger Centre  
 The Sanger Centre.

## FEATURES

source  
 1..642  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /map="20"  
 /clone\_11b="pool\_AK\_11b\_v\_SPD"  
 /note="Organ: breast; Vector: pZeo-1; Site 1: SpH1;  
 Site 2: SpH1; Ductal carcinoma in situ, high-grade, comedo  
 'from 41 yo female. Library constructed in the laboratory  
 of K. Polyak, M.D., Ph.D. (Dana-Farber Cancer Institute)."  
 from a pool of up to 100,000 cDNA clones derived from  
 pool\_AK\_11b\_v\_SPD cDNA library. Further information can be found at  
 http://www.sanger.ac.uk/teams/teams5/.

## BASE COUNT

140 a 164 c 181 g 157 t

Query Match 24.6%; Score 630.8; DB 9; Length 642;  
 Best Local Similarity 98.9%; Pred. No. 2, 9e-88;  
 Matches 635; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 1031 TACCTGAGTCTGACAGAGACACCACTCACTGAGAGAGTGGAGAGTGGCGGATGACATCTGCTCCCTCACTCAAGAGGCGCTGT 1090  
 Db 642 TACCTGAGTCTGACAGAGACACCACTCACTGAGAGAGTGGAGAGTGGCGGATGACATCTGCTCCCTCACTCAAGAGGCGCTGT 583  
 Qy 1091 TTTTGAAGCTGACAGAGAGAGTGGAGAGTGGCGGATGACATCTGCTCCCTCACTCAAGAGGCGCTGT 1150  
 Db 582 TTTTGAAGCTGACAGAGAGAGTGGAGAGTGGCGGATGACATCTGCTCCCTCACTCAAGAGGCGCTGT 523  
 Qy 1151 GCTTCAATCACTGAGAGAGTGGAGAGTGGCGGATGACATCTGCTCCCTCACTCAAGAGGCGCTGT 1210  
 Db 522 GCTTCAATCACTGAGAGAGTGGAGAGTGGCGGATGACATCTGCTCCCTCACTCAAGAGGCGCTGT 463  
 Qy 1211 GAGGCGAGAGAGAGAGAGTGGAGAGTGGCGGATGACATCTGCTCCCTCACTCAAGAGGCGCTGT 1270  
 Db 462 GAGGCGAGAGAGAGAGAGTGGAGAGTGGCGGATGACATCTGCTCCCTCACTCAAGAGGCGCTGT 403  
 Qy 1271 CCAAGGCAAGCTGACAGAGAGTGGAGAGTGGCGGATGACATCTGCTCCCTCACTCAAGAGGCGCTGT 1330  
 Db 402 CCAAGGCAAGCTGACAGAGAGTGGAGAGTGGCGGATGACATCTGCTCCCTCACTCAAGAGGCGCTGT 343  
 Qy 1331 CCACTGAGAGTGGAGAGTGGAGAGTGGCGGATGACATCTGCTCCCTCACTCAAGAGGCGCTGT 1390  
 Db 342 CCACTGAGAGTGGAGAGTGGAGAGTGGCGGATGACATCTGCTCCCTCACTCAAGAGGCGCTGT 283  
 Qy 1391 GCCATGATCCCACTGAGAGAGTGGAGAGTGGCGGATGACATCTGCTCCCTCACTCAAGAGGCGCTGT 1450  
 Db 282 GCCATGATCCCACTGAGAGAGTGGAGAGTGGCGGATGACATCTGCTCCCTCACTCAAGAGGCGCTGT 223  
 Qy 1451 TTTCAAAAGAGATAGAGTGGAGAGTGGCGGATGACATCTGCTCCCTCACTCAAGAGGCGCTGT 1510  
 Db 222 TTTCAAAAGAGATAGAGTGGAGAGTGGCGGATGACATCTGCTCCCTCACTCAAGAGGCGCTGT 163  
 Qy 1511 TTTCAAGTCACTGAGAGAGTGGAGAGTGGCGGATGACATCTGCTCCCTCACTCAAGAGGCGCTGT 1570  
 Db 162 TTTCAAGTCACTGAGAGAGTGGAGAGTGGCGGATGACATCTGCTCCCTCACTCAAGAGGCGCTGT 103  
 Qy 1571 AGGCGCGACCCAGAGTGAAGAACCCCTAGAGTCAACGAGAAAGTCAATTTTCAGAA 1630  
 Db 102 AGGCGCGACCCAGAGTGAAGAACCCCTAGAGTCAACGAGAAAGTCAATTTTCAGAA 43  
 Qy 1631 ATCTCAAGTCTGTTGAGAGAGTGGAGAGTGGCGGATGACATCTGCTCCCTCACTCAAGAGGCGCTGT 1672  
 Db 42 ATCTCAAGTCTGTTGAGAGAGTGGAGAGTGGCGGATGACATCTGCTCCCTCACTCAAGAGGCGCTGT 1



```

RESULT 11
AL844312/c 619 bp mRNA linear EST 30-JUL-2002
LOCUS
DEFINITION AL844312 pool_YT_11b_v_SPD Homo sapiens cDNA, mRNA sequence.
ACCESSION AL844312
VERSION AL844312.1 GI:22019094
KEYWORDS
SOURCE EST.
ORGANISM human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 619)
Ashcroft,K., Bethel,G., Bye,J.M., Howell,G.R., Huckle,E.J. and
Sheridan,E.
TITLE Homo sapiens EST sequence
JOURNAL Unpublished (2002)
COMMENT Contact: The Sanger Centre
The Sanger Centre
Wellcome Trust Genome Campus, Hinxton, Cambridgeshire, CB10 1SA, UK
Email: humquery@sanger.ac.uk
Sanger Centre name: scdd10818.400489A
Homo sapiens EST sequence. This sequence was generated as part of
The Wellcome Trust Sanger Institute program to identify and
annotate genes in the human genome. Incomplete or unconfirmed genes
are experimentally analysed using a variety of cDNA library
resources. This sequence was obtained from a PCR product generated
from a pool of up to 100,000 cDNA clones derived from
pool_YT_11b_v_SPD cDNA library. Further information can be found at
http://www.sanger.ac.uk/Teams/Team69/.
Location/Qualifiers
1..619
/organism="Homo sapiens"
/db_xref="taxon:9606"
/map="20"
/clone_lib="pool_YT_11b_v_SPD"
/note="Organ: breast; Vector: pZero-1; Site: 1: Spl1;
Site_2: Spl1; Ductal carcinoma in situ, high-grade, comedo
, from 41 yo female. Library constructed in the laboratory
of K. Polyak, M.D., Ph.D. (Dana-Farber Cancer Institute)."
```

BASE COUNT 133 a 159 c 177 g 150 t

ORIGIN

Query Match 24.0%; Score 615.8; DB 9; Length 619;  
 Best Local Similarity 99.7%; Pred. No. 6.2e-86;  
 Matches 617; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

QY 1036 GTGACTGTGAGAGACACCACTGAGAAAGAGCTGACAGCTCCCTCTGTTTCT 1095
DB 619 GTGACTGTGAGAGACACCACTGAGAAAGAGCTGACAGCTCCCTCTGTTTCT 560
QY 1096 GAAGCTGCACAGGAGAGAGCTCTCTAGTGAAGGCTCCGGAGTCCCTCAGCTTC 1155
DB 559 GAAGCTGCACAGGAGAGAGCTCTCTAGTGAAGGCTCCGGAGTCCCTCAGCTTC 500
QY 1156 TACATCAGCTTGATGACGAGGCTGTCTCTTTGATGATGCTTGAAGGCCCAAGAGAGGC 1215
DB 499 TACATCAGCTTGATGACGAGGCTGTCTCTTTGATGATGCTTGAAGGCCCAAGAGAGGC 440
QY 1216 CAAAAGGAAACCAAGGCTGCACACCTAGAACCCCAATTGACCTCTGGGACCCCGCA 1275
DB 439 CAAAAGGAAACCAAGGCTGCACACCTAGAACCCCAATTGACCTCTGGGACCCCGCA 380
QY 1276 GGCAGGCTGTGCACTCAGGAGGAGGAGGAGGAGACACAGAGGTGATAGAGTCCAC 1315
DB 379 GGCAGGCTGTGCACTCAGGAGGAGGAGGAGGAGGAGACACAGAGGTGATAGAGTCCAC 320
QY 1336 TGTACCTGTGCTCTTCTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1395
DB 319 TGTACCTGTGCTCTTCTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 260
QY 1396 GATCCACCTGCGAAGCTTCTAGTGAAGTGAAGAAAGTGGAGACAGAGGCTTCA 1455
DB 259 GATCCACCTGCGAAGCTTCTAGTGAAGTGAAGAAAGTGGAGACAGAGGCTTCA 200
```

---

```

QY 1456 AAAAGAGATTAAGCTCTCTGAGGAGGAGTCTGACCTAGTATGTTGATTGGGTTTCCA 1515
DB 199 AAAAGAGATTAAGCTCTCTGAGGAGGAGTCTGACCTAGTATGTTGATTGGGTTTCCA 140
QY 1516 GTACCATCTGAGATGCGCTGCTGTTGAGGCCCATTCATCCGCCACATTAACGAGGCC 1575
DB 139 GTACCATCTGAGATGCGCTGCTGTTGAGGCCCATTCATCCGCCACATTAACGAGGCC 80
QY 1576 CCACCCCAAGAGTGAAGAAACCCCTTGAAGTCAACGAGAAAGTCAATTTTCAGAAATCTA 1635
DB 79 CCACCCCAAGAGTGAAGAAACCCCTTGAAGTCAACGAGAAAGTCAATTTTCAGAAATCTA 20
QY 1636 CAAGTCTGTTGAGACCCAC 1654
DB 19 CAAGTCTGTTGAGACCCAC 1
```

RESULT 12

```

AL844307 614 bp mRNA linear EST 30-JUL-2002
LOCUS
DEFINITION AL844307 pool_AK_11b_v_SPD Homo sapiens cDNA, mRNA sequence.
ACCESSION AL844307
VERSION AL844307.1 GI:22019089
KEYWORDS
SOURCE EST.
ORGANISM human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 614)
Ashcroft,K., Bethel,G., Bye,J.M., Howell,G.R., Huckle,E.J. and
Sheridan,E.
TITLE Homo sapiens EST sequence
JOURNAL Unpublished (2002)
COMMENT Contact: The Sanger Centre
The Sanger Centre
Wellcome Trust Genome Campus, Hinxton, Cambridgeshire, CB10 1SA, UK
Email: humquery@sanger.ac.uk
Sanger Centre name: scdd10818.154136A
Homo sapiens EST sequence. This sequence was generated as part of
The Wellcome Trust Sanger Institute program to identify and
annotate genes in the human genome. Incomplete or unconfirmed genes
are experimentally analysed using a variety of cDNA library
resources. This sequence was obtained from a PCR product generated
from a pool of up to 100,000 cDNA clones derived from
pool_AK_11b_v_SPD cDNA library. Further information can be found at
http://www.sanger.ac.uk/Teams/Team69/.
Location/Qualifiers
1..614
/organism="Homo sapiens"
/db_xref="taxon:9606"
/map="20"
/clone_lib="pool_AK_11b_v_SPD"
/note="Organ: breast; Vector: pZero-1; Site: 1: Spl1;
Site_2: Spl1; Ductal carcinoma in situ, high-grade, comedo
, from 41 yo female. Library constructed in the laboratory
of K. Polyak, M.D., Ph.D. (Dana-Farber Cancer Institute)."
```

BASE COUNT 134 a 188 c 164 g 128 t

ORIGIN

Query Match 23.7%; Score 609.2; DB 9; Length 614;  
 Best Local Similarity 99.5%; Pred. No. 6.5e-85;  
 Matches 611; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

QY 882 CTGCGTGTACATTCACCGGCTCCTGCTCCCTGACCTGAGGCTCTGAGACATTA 941
DB 1 CTGCGTGTACATTCACCGGCTCCTGCTCCCTGACCTGAGGCTCTGAGACATTA 60
QY 942 CTGAGCTGGGAGTGAATCTGCTGCTTCTGCTTCTGAGAGGCTCTGAGAGGCT 1001
DB 61 CTGAGCTGGGAGTGAATCTGCTGCTTCTGCTTCTGAGAGGCTCTGAGAGGCT 120
QY 1002 TGGCGCGCTCCTGAGCAAGATTAACCTTACTGTGATCTGAGAGAGACCACTCA 1061
```

Db 121 TGGCCCGCTCCCTGGCAAGATATACCTTACCTGATGCTGACAGACACACCTCAA 180  
 Qy 1062 CTGGAAAGAGCTGACAGCTCCCTCTGTTTCTGAAAGCTGACAGAGGAGAGCTCT 1121  
 Db 181 CTGGAAAGAGCTGACAGCTCCCTCTGTTTCTGAAAGCTGACAGAGGAGAGCTCT 240  
 Qy 1122 TCTCAGTAGAGGCTCTCCGGAGTCCCTCACTTCACTCACTGAAATGACAGAGCT 1181  
 Db 241 TCTCAGTAGAGGCTCTCCGGAGTCCCTCACTTCACTCACTGAAATGACAGAGCT 300  
 Qy 1182 CTCTTTGATGATGCTTACAGCCCAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1241  
 Db 301 CTCTTTGATGATGCTTACAGCCCAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGG 360  
 Qy 1242 TAGAACCCCAATTACCTCTCTGGGACCCCAAGAGGAGGAGGAGGAGGAGGAGG 1301  
 Db 361 TAGAACCCCAATTACCTCTCTGGGACCCCAAGAGGAGGAGGAGGAGGAGGAGGAG 420  
 Qy 1302 GGGTGGACACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1361  
 Db 421 GGGTGGACACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 480  
 Qy 1362 CCTTGAAGAGTACCTTCTCTTCCAGTGCATGATCCCACTGAGCTCTTCTCTT 1421  
 Db 481 CCTTGAAGAGTACCTTCTCTTCCAGTGCATGATCCCACTGAGCTCTTCTCTT 540  
 Qy 1422 GTGCAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1481  
 Db 541 GTGCAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 600  
 Qy 1482 CTGACCTAGTTACT 1495  
 Db 601 CTGACCTAGTTACT 614

RESULT 13  
 AL844310/c 611 bp mRNA linear EST 30-JUL-2002  
 LOCUS AL844310 pool\_FLU\_1lb\_v\_SPC Homo sapiens cDNA, mRNA sequence.  
 DEFINITION AL844310 pool\_FLU\_1lb\_v\_SPC Homo sapiens cDNA, mRNA sequence.  
 ACCESSION AL844310  
 VERSION AL844310.1 GI:22019092  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 611)  
 AUTHORS Aehroldt, K., Bechtel, G., Bye, J.M., Howell, G.R., Huckle, E.J. and Sheridan, E.  
 TITLE Homo sapiens EST sequence  
 JOURNAL Unpublished (2002)  
 COMMENT Contact: The Sanger Centre  
 The Sanger Centre  
 Wellcome Trust Genome Campus, Hinxton, Cambridgeshire, CB10 1SA, UK  
 Email: humquery@sanger.ac.uk  
 Sanger Centre name: scdcl0817.400489A  
 Homo sapiens EST sequence. This sequence was generated as part of  
 The Wellcome Trust Sanger Institute program to identify and  
 annotate genes in the human genome. Incomplete or unconfirmed genes  
 are experimentally analysed using a variety of cDNA library  
 resources. This sequence was obtained from a PCR product generated  
 from a pool of up to 100,000 cDNA clones derived from  
 pool FLU\_1lb v\_SPC cDNA library. Further information can be found  
 at <http://www.sanger.ac.uk/Teams/Team9/>.  
 Location/Qualifiers  
 1..611  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /map="20"  
 /clone\_1lb="pool\_FLU\_1lb v\_SPC"  
 /notes="Organ: breast; Vector: pBEO-1; Site 1: SphI;  
 Site 2: SphI; Ductal carcinoma in situ, high-grade, comedo

BASE COUNT 132 a 157 c 173 g 149 t  
 ORIGIN  
 Query Match 23.4%; Score 601.4; DB 9; Length 611;  
 Best Local Similarity 99.0%; Pred. No. 1,1e-83;  
 Matches 605; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Db 1061 ACTGAAAGAGCTGACAGCTCCCTCTGTTTCTGAAAGCTGACAGAGGAGAGCTC 1120  
 Db 611 ACTGAAAGAGCTGACAGCTCCCTCTGTTTCTGAAAGCTGACAGAGGAGAGCTC 552  
 Qy 1121 TTCTCAGTAGAGGCTCTCCGGAGTCCCTCACTTCACTCACTGAAATGACAGAGCT 1180  
 Db 551 TTCTCAGTAGAGGCTCTCCGGAGTCCCTCACTTCACTCACTGAAATGACAGAGCT 492  
 Qy 1181 TCTCTTGAATGATGCTTACAGCCCAAGAGAGGAGGAGGAGGAGGAGGAGGAGG 1240  
 Db 491 TCTCTTGAATGATGCTTACAGCCCAAGAGAGGAGGAGGAGGAGGAGGAGGAGG 432  
 Qy 1241 CTGAACCCCAATTACCTCTCTGGGACCCCAAGAGGAGGAGGAGGAGGAGGAGG 1300  
 Db 431 CTGAACCCCAATTACCTCTCTGGGACCCCAAGAGGAGGAGGAGGAGGAGGAGG 372  
 Qy 1301 AGGTTGGAGACACAGAGTGCATTAAGGATCCCACTGTAACCTTCTCTCTT 1360  
 Db 371 AGGTTGGAGACACAGAGTGCATTAAGGATCCCACTGTAACCTTCTCTCTT 312  
 Qy 1361 GCCCTTGAAGTACCTTCTCTTCCAGTGCATGATCCCACTGAGCTCTTCTCTT 1420  
 Db 311 GCCCTTGAAGTACCTTCTCTTCCAGTGCATGATCCCACTGAGCTCTTCTCTT 252  
 Qy 1421 AGTGCAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1480  
 Db 251 AGTGCAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 192  
 Qy 1481 TCTGACCTAGTTACTCTTGAATGTTGGGATTTCCAGTACCATCTGATGCTCTGTT 1540  
 Db 191 TCTGACCTAGTTACTCTTGAATGTTGGGATTTCCAGTACCATCTGATGCTCTGTT 132  
 Qy 1541 GAGCCCCCTTCTACATCCCAACCAATTAACAGAGGAGGAGGAGGAGGAGGAGG 1600  
 Db 131 GAGCCCCCTTCTACATCCCAACCAATTAACAGAGGAGGAGGAGGAGGAGGAGG 72  
 Qy 1601 TAGAGTCAACGAGAAATGATTTTCAAGAAATCTACAGTCTGTTGAGACACCAT 1660  
 Db 71 TAGAGTCAACGAGAAATGATTTTCAAGAAATCTACAGTCTGTTGAGAGGAGGAGG 12  
 Qy 1661 ACCTCAGAAAG 1671  
 Db 11 ACCTCAGAAAG 1

RESULT 14  
 B0053486 878 bp mRNA linear EST 29-MAR-2002  
 LOCUS B0053486  
 DEFINITION AGENCOUR\_6822017 NIH\_MGC\_106 Homo sapiens cDNA clone IMAGE:5935253  
 5', mRNA sequence.  
 ACCESSION B0053486  
 VERSION B0053486.1 GI:19812826  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 878)  
 AUTHORS NIH-MGC <http://mgi.nhl.gov/>  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: [cgabs-remail.nih.gov](mailto:cgabs-remail.nih.gov)  
 Tissue Procurement: Dr. Daniel McVicar, DBS/NCI



Tue Apr 1 06:01:26 2003

us-09-988-971-1.rst

Page 12

|    |      |   |      |
|----|------|---|------|
| Qy | 1180 | GTCTTTTGGATAGTATGCTTAAAGCCCAAAAGAGAGCCAAAAGGAGAAACCAAGCTGTCA  | 1239 |
| Db | 301  | GTCTCTTTGGATATCTTAAGCCCAAAAGAGAGCCAAAAGGAGAAACCAAGGCTGTCA     | 360  |
| Qy | 1240 | CCTTAAGACCCCAATTACAGCTTCTTGGGCAACCCAGAGGCAAGCTGTGTCACTCAAGGAG | 1239 |
| Db | 361  | CCTAAGACCCCAATTACAGCTTCTTGGGCAACCCAAAAGCAAACTGTGTCACTCAAGGAG  | 420  |
| Qy | 1300 | GAGGTGGGACACAGAGGTGCATTAAGGTGCCACATGTATCCCTTCTCTCTCTT         | 1359 |
| Db | 421  | GAGGTGGGACACAAAGGTGCATTAAGGTGCCACATGTATCCCTTCTCTCTCTCTT       | 480  |
| Qy | 1360 | AGCCTTTGAAATGACCTATCTTCTTCAAGTGCATATGCCACCTGGGACCTCTATGTC     | 1419 |
| Db | 481  | AGCCTTTGAAATGACCTATCTTCTTCAAGTGCATATGCCACCTGGGACCTCTATGTC     | 540  |
| Qy | 1420 | GAGTGCAGAGAGGTGGGACCAAGGCGCAAGGGTTCCAAAAAGAGATTAAGCTTCTGGGG   | 1479 |
| Db | 541  | GAGTGCAGAGAGGTGGGACCAAGGCGCAAGGTTCCAAAAAGAGATTAAGCTTCTGGGG    | 600  |
| Qy | 1480 | GTCTGA 1495   |      |
| Db | 601  | GTCTGA 606  |      |

Search completed: March 30, 2003, 02:49:10  
Job time : 5065.55 BECS